

Tracking Deforming Objects by Filtering and Prediction in the Space of Curves

Ganesh Sundaramoorthi* Andrea Mennucci† Stefano Soatto* Anthony Yezzi‡

* Computer Science Department, University of California, Los Angeles - CA, USA ({ganeshs, soatto}@ucla.edu)

† Scuola Normale Superiore, Pisa, Italy (a.mennucci@sns.it)

‡ Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta - GA, USA (ayezzi@ece.gatech.edu)

Abstract—We propose a dynamical model-based approach for tracking the *shape* and *deformation* of highly deforming objects from time-varying imagery. Previous works have assumed that the object deformation is *smooth*, which is realistic for the tracking problem, but most have restricted the deformation to belong to a finite-dimensional group, such as affine motions, or to finitely-parameterized models. This, however, limits the accuracy of the tracking scheme. We exploit the smoothness assumption implicit in previous work, but we lift the restriction to finite-dimensional motions/deformations. To do so, we derive analytical tools to define a dynamical model on the (infinite-dimensional) space of curves. To demonstrate the application of these ideas to object tracking, we construct a simple dynamical model on shapes, which is a first-order approximation to any dynamical system. We then derive an associated nonlinear filter that estimates and predicts the shape and deformation of a object from image measurements.

I. INTRODUCTION

Our goal in this paper is to track highly deforming objects from time-varying two-dimensional images. We are interested in tracking the precise *shape* of the deforming object by predicting and extrapolating its deformation. The class of objects we focus on can be described by non-self-intersecting closed planar curves, as customary in the literature of *active contours* [12], although most of that literature focused on static imagery. Extensions to temporal data typically involves two steps. One is the collection of local statistics from a single image (e.g. intensity histograms, spatial and temporal regularized derivatives etc.). There is a vast literature on image domain partitioning based on such local statistics using a variety of criteria and their associated energy functionals, from intensity-based segmentation [3], [17], [4] to piecewise parametric optical flow estimation [7], [20]. The other step is incorporating a model of the temporal variation of the deforming object into the tracking algorithm. The simplest way to extend the active contour methodology to time-varying imagery is to use the contour estimated at time “ t ” as initialization for the same gradient-based optimization at time “ $t+1$ ” [4]. This approach implicitly assumes trivial dynamics (“constant position plus perturbation”), so its prediction would trail an object moving with constant speed with a constant error. Better dynamics (“constant velocity plus perturbation”) have been developed both for parametric [2], [24], [10] and geometric [21], [18], [11] active contour models, the latter implemented using level set methods [19]. While these methods can more accurately predict the (*affine*) *motion* of the object, their deformation model remains trivial

(no deformation on average): So, the prediction of the motion of a jelly fish would extrapolate its affine trajectory (position, orientation, scale and skew) but “freeze” its shape to the last observation. Thus the dynamical model – and therefore the predictive ability of the tracking scheme – is restricted to the finite-dimensional portion of the actual object deformation.

Recent work has moved beyond the assumption of affine motion [6], [25]. In [6], the motion/deformation is described by a linear autoregressive model defined on combinations of distance functions given as a training set. The applicability of this method is therefore restricted by the availability of training data for every particular object class and its associated deformations. In [25], the authors use a small time-varying basis, which is finite-dimensional but beyond affine, to dynamically model local deformations of the contour.

In this paper, we define a dynamical model *directly* on the infinite-dimensional space of (closed, simple, planar) curves to model the deformation of the object of interest. The study of shapes as points on an infinite dimensional space has been the subject of considerable interest recently [5], [16], [1], [8], [28]. In the present work, we construct a Riemannian structure on the space of curves using a geometric-type Sobolev metric [27], [15], [23]. This is useful for tracking because it favors *smooth* motions of the curve without restricting its deformation. To illustrate these ideas, we construct a simple constant (infinite-dimensional) velocity model of the contour and then derive an associated filter that predicts and estimates the contour and its deformation based on local image statistics. For simplicity, in this work we consider simple intensity statistics, but local spatio-temporal filters can be used as well.

II. GEOMETRY IN THE SPACE OF CURVES

We define the geometry on the space of “shapes,” defined from now on as the set of simple, closed planar curves, or “contours”. This is the foundation for constructing dynamical systems and filtering strategies on deforming shapes.

A. The Space of Curves

We define the space of plane immersed curves as

$$M = \{c \in C^\infty(\mathbb{S}^1, \mathbb{R}^2) : |c'(\theta)| \neq 0 \forall \theta \in \mathbb{S}^1\}; \quad (\text{II.1})$$

we are interested in *geometric* curves, i.e., considered up to reparameterizations; thus, we define the quotient space

$$B = M/\text{Diff}(\mathbb{S}^1) \quad (\text{II.2})$$

where $\text{Diff}(\mathbb{S}^1)$ denotes the group of diffeomorphisms of \mathbb{S}^1 , the circle. It can be shown that M and B are manifolds [14], and the tangent space $T_c M$ at c of M is the set of *deformations* $h \in T_c M$, which are vector fields on c , i.e., $h : \mathbb{S}^1 \rightarrow \mathbb{R}^2$. In order to define a (Riemannian) metric, we must specify an inner product $\langle \cdot, \cdot \rangle$ on $T_c M$, which we do in the next section, so that M and B become *Riemannian manifolds*. To define the *horizontal space* of $T_c M$, i.e., the tangent space to B at c , $T_{[c]} B$, we first define the *vertical space* as

$$V_c M = \{h \in T_c M : h = \beta c_\theta, \beta : \mathbb{S}^1 \rightarrow \mathbb{R}\}. \quad (\text{II.3})$$

This is the set of deformations of c that do not change the geometry of the curve c . Given an inner product on $T_c M$, we may define the *horizontal space* as

$$W_c M \doteq T_{[c]} B = (V_c M)^\perp = \{h \in T_c M : \langle h, k \rangle = 0, \forall k \in V_c M\}. \quad (\text{II.4})$$

Assuming that the inner product on $T_c M$ defines a Riemannian metric, we are able to define distances on B , geodesics (i.e., shortest paths), the exponential map, and parallel transport. All these operations will be essential to define dynamical models in the infinite-dimensional space of curves, as described in the previous section. We defer the computation of these operations to the next section, and restrict ourselves to their definition in this section. Given two curves¹ $[c_0]$ and $[c_1]$, we define

Definition II.5: The distance, $d : B \times B \rightarrow \mathbb{R}^+$, on B or M is

$$d([c_0], [c_1]) = \inf_{\phi \in \text{Diff}(\mathbb{S}^1)} \inf_{\gamma \in \Gamma(c_0, c_1 \circ \phi)} \text{Len}(\gamma) \quad (\text{II.5.*})$$

where $\Gamma(c_0, c_1) = \{\gamma : [0, 1] \rightarrow M : \gamma(0) = c_0, \gamma(1) = c_1\}$,

$$\text{Len}(\gamma) = \int_0^1 \|\dot{\gamma}(t)\|_{\gamma(t)} dt, \quad (\text{II.5.†})$$

$\dot{\gamma}(t) \in T_{\gamma(t)} M$ is the velocity or deformation of $\gamma(t)$, and $\|\cdot\|_{\gamma(t)}$ is the norm on $T_{\gamma(t)} M$.

It can be shown that the infima above must be attained by a path γ^* that satisfies $\dot{\gamma}^*(t) \in W_{\gamma(t)} M$. We define

Definition II.6: A *geodesic* between $[c_0]$ and $[c_1]$ is a path γ^* that attains the infimum in (II.5.*). Equivalently, up to a re-parameterization of the path γ^* , γ^* solves

$$\inf_{\phi \in \text{Diff}(\mathbb{S}^1)} \inf_{\gamma \in \Gamma(c_0, c_1 \circ \phi)} \mathbb{E}(\gamma) \quad (\text{II.6.*})$$

where

$$\mathbb{E}(\gamma) = \int_0^1 \|\dot{\gamma}(t)\|_{\gamma(t)}^2 dt \quad (\text{II.6.†})$$

is the energy of the path γ .

We may now define the *exponential map* as

Definition II.7: The *exponential map*, $\exp : TB \rightarrow B$, where TB is the tangent bundle of B , is

$$\exp_{[c]}(h) = \gamma(1),$$

¹The notation $[c]$ indicates that a curve represents an equivalence class under the equivalence relation (II.2).

where $\gamma : [0, 1] \rightarrow B$ is the geodesic with $\dot{\gamma}(0) = h \in T_{[c]} B$.

Definition II.8: The inverse of the exponential map, which we call the *logarithm*, $\log : B \times B \rightarrow TB$ is

$$\log([c_0], [c_1]) = \dot{\gamma}(0) \in T_{[c_0]} B$$

where $\gamma : [0, 1] \rightarrow B$ satisfies $\gamma(0) = [c_0]$, $\gamma(1) = [c_1]$ and is the geodesic between $[c_0]$ and $[c_1]$.

Finally, we define *parallel transport* as

Definition II.9: The *parallel transport*, $P_{\gamma, t_0, t_1} : T_{\gamma(t_0)} B \rightarrow T_{\gamma(t_1)} B$, along a path $\gamma : [0, 1] \rightarrow B$ from $t_0 \in [0, 1]$ to $t_1 \in [0, 1]$ of the tangent vector $h \in T_{\gamma(t_0)} B$ is

$$P_{\gamma, t_0, t_1}(h) = V(t_1)$$

where the vector field, $V(t) \in T_{\gamma(t)} B$ is such that

$$\nabla_{\gamma'(t)} V(t) = 0, \quad V(0) = h$$

where $\nabla_{\gamma'}$ is the covariant derivative along γ .

B. A Geometric Sobolev-Type Metric

We now define a Riemannian metric H on the tangent space $T_c M$. First, for $c \in M$ and $h \in T_c M$, we define the following decomposition:

$$h = h^t + h^l \Pi_1(c) + h^d \quad (\text{II.10})$$

where h^t is the component of h that translates the centroid of c , $h^l \Pi_1(c)$ is the component of h that changes the scale (length) of c , and $h^d = h - h^t - h^l \Pi_1(c)$ is the component of h that deforms c without scale or translation. The components h^t and h^l of h are defined as

$$h^t = d\bar{c} \cdot h = \oint p(h) ds \in \mathbb{R}^2 \quad (\text{II.11})$$

$$h^l = dL(c) \cdot h = \oint h \cdot D_s^2 c ds \in \mathbb{R} \quad (\text{II.12})$$

where s is the arclength parameter of c and

$$\oint \cdot ds \doteq \frac{1}{L(c)} \int_c \cdot ds \quad (L(c) = \text{length of } c)$$

$$\Pi_1(c) \doteq c - \bar{c} \quad (\bar{c} = \text{centroid of } c)$$

$$p(h) \doteq h - (h \cdot D_s c) D_s c - (h \cdot D_s^2 c) \Pi_1(c),$$

If $h, k \in T_c M$, then we define the Riemannian metric as

$$\langle h, k \rangle_{H_c} \doteq h^t \cdot k^t + h^l k^l + \oint D_s h^d \cdot D_s k^d ds, \quad (\text{II.13})$$

where the first two products are the Euclidean dot products, and the last term is a normalized geometric Sobolev metric. Therefore, in this metric, centroid translations, scale changes and deformations of the shape are orthogonal. We choose this metric for the following reasons:

- 1) Sobolev-type metrics favor *smooth* but otherwise unrestricted infinite-dimensional deformations, a desirable property for tracking a highly deforming object.
- 2) Sobolev-type metrics have proven useful in framewise image segmentation for visual tracking because of their coarse-to-fine evolution behavior [23].
- 3) It will be shown in the next section that geodesics in

this metric (i.e., the optimization problem in (II.5.*)) can be computed efficiently.

C. Computing Geodesics, exp, and log

Let $C : [0, 2\pi] \times [0, 1] \rightarrow \mathbb{R}^2$, $(\theta, t) \rightarrow C(\theta, t)$ denote a time varying family of closed curves (i.e., a homotopy) corresponding to some path $\gamma : [0, 1] \rightarrow M$, i.e., $C(\theta, t) = \gamma(t)(\theta)$. Then we have that

$$\|\partial_t C\|_H^2 = |\partial_t \bar{C}|^2 + (\partial_t(\log L(C)))^2 + \int |D_s(\partial_t C)^d|^2 ds \quad (\text{II.14})$$

Using the above fact and some manipulation one can show that geodesics in this metric are invariant to scale and translations:

Proposition II.15 (Invariance of geodesics.): Let $c_0, c_1 \in M$ and let $\tilde{c}_1 = v + e^\lambda(c_1 - \bar{c}_1)$ be a scaling and translation of c_1 . Suppose that C is a homotopy connecting c_0 to c_1 , and let $\tilde{C} = tv + e^{t\lambda}(C - \bar{C})$ be a homotopy connecting c_0 to \tilde{c}_1 ; then

$$\mathbb{E}(C) = \mathbb{E}(\tilde{C}) + \text{const}$$

where the “constant term” depends only the end curves c_0, c_1 and on v, λ . As a corollary we obtain that C is a geodesic connecting c_0 to c_1 if and only if \tilde{C} is a geodesic connecting c_0 to \tilde{c}_1 .

The conservation of momenta and (II.14) imply that

Proposition II.16 (Momenta): • [translation] $\partial_t \bar{C}$ is constant along geodesics;

- [scaling] $\partial_t(\log L(C))$ is constant along geodesics;
- [rotation] $\langle AC, \partial_t C \rangle_H$ is constant along geodesics for any antisymmetric matrix A ;
- [reparametrization] $\langle f D_s C, \partial_t C \rangle_H$ is constant along geodesics for any $f : \mathbb{S}^1 \rightarrow \mathbb{R}$.

This in particular means that along any geodesic C ,

$$\begin{aligned} \bar{C} &= (1-t)\bar{c}_0 + t\bar{c}_1 \\ \log L(C) &= (1-t)\log L(c_0) + t\log L(c_1). \end{aligned} \quad (\text{II.17})$$

The previous result along with Proposition II.15 implies that to compute a geodesic in M between c_0 and c_1 , we apply the following procedure:

1) Define

$$\tilde{c}_0 \doteq \frac{c_0 - \bar{c}_0}{L(c_0)}, \quad \tilde{c}_1 \doteq \frac{c_1 - \bar{c}_1}{L(c_1)}$$

- 2) Compute a geodesic \tilde{C} between \tilde{c}_0 and \tilde{c}_1 in the space M_d of unit length curves with centroid at the origin
- 3) Rebuild the geodesic in M :

$$C(t, \cdot) = L^{1-t}(c_0).L^t(c_1).\tilde{C}(t, \cdot) + (1-t)\bar{c}_0 + t\bar{c}_1.$$

We have therefore reduced the problem to computing geodesics in the space M_d of unit length curves with centroid at the origin. To do this, we exploit the map Φ introduced by Younes et al. in [27], [28]: let $e, f \in V_{od} \doteq \{f : \mathbb{S}^1 \rightarrow \mathbb{R} | f(0) = -f(2\pi)\}$ then

$$c(\xi) = \Phi(e, f)(\xi) \doteq \frac{1}{2} \int_0^\xi (e + if)^2(\theta) d\theta \quad (\text{II.19})$$

where i denotes the imaginary unit. The inverse of Φ exists, and is known as the *square-root lifting*. Note that for c above to be a closed curve and of unit length, we must have that (e, f) belong to

$$\begin{aligned} \text{St}(2, V_{od}) = \\ \{(e, f) \in V_{od} : \|e\|_{\mathbb{L}^2} = \|f\|_{\mathbb{L}^2} = 1, \langle e, f \rangle_{\mathbb{L}^2} = 0\} \end{aligned}$$

where the above \mathbb{L}^2 norms and inner product are the standard ones on $\mathbb{L}^2([0, 2\pi])$. The above set, $\text{St}(2, V_{od})$, is known as a *Stiefel* manifold, which is classically defined as the set of all p ($1 \leq p \leq n$) orthonormal vectors in \mathbb{R}^n . It is shown in [28] that Φ is an isometry between $\text{St}(2, V_{od})$ and M_d endowed with the Sobolev metric:

Theorem II.20 (2.2 in [28]): Let $c^d \in M_d$, $h^d \in T_{c^d} M_d$ and $(e, f) \in \text{St}(2, V_{od})$, $(\delta e, \delta f) \in T_{(e,f)} \text{St}(2, V_{od})$ be the corresponding Stiefel representations. Then

$$\int_{c^d} |D_s h^d|^2 ds = \int_0^{2\pi} (\delta e)^2 + (\delta f)^2 d\theta,$$

that is, the Sobolev-type Riemannian metric in M_d is mapped to the flat metric \mathbb{L}^2 in $\text{St}(2, V_{od})$.

Therefore, the theorem above allows us to compute geodesics in M_d by computing the corresponding geodesics in $\text{St}(2, V_{od})$ and then mapping them back to M_d via the map Φ .

Geodesics in Stiefel manifolds, $\text{St}(p, \mathbb{R}^n)$, are known to have closed form solutions as demonstrated by [9]:

Proposition II.21 ([9]): Let $Y : [0, 1] \rightarrow \text{St}(p, \mathbb{R}^n)$ be a path, then the geodesic equation (when the Euclidean metric is used) is

$$\ddot{Y} + Y(\dot{Y}^T \dot{Y}) = 0. \quad (\text{II.21.*})$$

The solution is

$$(Y(t)e^{At}, \dot{Y}(t)e^{At}) = (Y(0), \dot{Y}(0)) \exp t \begin{pmatrix} A & -S \\ I & A \end{pmatrix} \quad (\text{II.21.†})$$

where $A = Y^T(0)\dot{Y}(0)$, $S = \dot{Y}^T(0)\dot{Y}(0)$, and I is the $p \times p$ identity matrix.

The solution, while written for $\text{St}(p, \mathbb{R}^n)$, extends to $\text{St}(2, V_{od})$. Indeed, (II.21.†) shows that $Y(t)$ remains in the space spanned by $\{Y(0), \dot{Y}(0)\}$ for all t . Let $(e, f) \in \text{St}(2, V_{od})$ and $(\delta e, \delta f) \in T_{(e,f)} \text{St}(2, V_{od})$ be initial conditions of the geodesic in $\text{St}(2, V_{od})$. We may define an orthonormal basis $\mathcal{B} = \{e, f, \tilde{e}, \tilde{f}\} \subset \text{span}(\{e, f, \delta e, \delta f\})$. Let $a, b \in \mathbb{R}^4$ be the representations of $\delta e, \delta f$ relative to \mathcal{B} , then we may define $Y(0) = (e_1, e_2)$ where $e_1 = (1, 0, 0, 0)^T$ and $e_2 = (0, 1, 0, 0)$ and $\dot{Y}(0) = (a, b)$ and then apply (II.21.†) to obtain the geodesic $Y(t) = (Y_1(t), Y_2(t))$ in the coordinates of the basis \mathcal{B} . The geodesic in $\text{St}(2, V_{od})$ is then given by

$$\begin{aligned} e(t) &= Y_1^1(t)e + Y_1^2(t)f + Y_1^3(t)\tilde{e} + Y_1^4(t)\tilde{f} \\ f(t) &= Y_2^1(t)e + Y_2^2(t)f + Y_2^3(t)\tilde{e} + Y_2^4(t)\tilde{f} \end{aligned}$$

where Y_j^i denotes the i^{th} component of $Y_i \in \mathbb{R}^4$.

The formula (II.21.†) gives the geodesic as a function of the initial position and direction; this is the exponential map.

However, to compute geodesics between two curves or the logarithmic map, it is necessary to have a formula for Y in terms of the boundary conditions $Y(0)$ and $Y(1)$. We are not aware of such a formula at this time, and therefore, to compute the geodesic and logarithm between $Y_0 = (e_1, e_2)$ and $Y_1 \in \mathbb{R}^4 \times \mathbb{R}^4$, we minimize the energy

$$E(\lambda, v_1, v_2, v_3, v_4) = |Y(1) - Y_1|^2 \quad (\text{II.22})$$

where $Y(1)$ is given according to (II.21.†) with the initial condition $\dot{Y}(0) = ((0, \lambda, v_1, v_3)^T, (-\lambda, 0, v_2, v_4)^T)$. We minimize (II.22) by standard gradient descent in \mathbb{R}^5 , and the gradient of (II.22) can be computed using a formula for matrix exponentials in [13]. This can be computed efficiently using an FFT technique similar to a technique found in [26], where we refer the reader for details.

Up to this point, we have specified how to compute geodesics, the logarithm and exponential maps in M according to the metric H ; however, we are interested in these operations in the geometric space B . To do this, we note that geodesics in B (with the metric induced from M) correspond to geodesics in M provided they are horizontal, i.e., $\dot{\gamma}(t) \in W_{\gamma(t)}M, \forall t$. Equivalently, it is enough that $\dot{\gamma}(1) \in W_{\gamma(1)}M$ and γ be a geodesic in M for γ to be a geodesic in B . Thus, to compute a geodesic between $[c_0], [c_1] \in B$, we iterate the following steps

- 1) Compute the geodesic, $(e_{\phi_k}(t), f_{\phi_k}(t)), t \in [0, 1]$, in M between (e_0, f_0) and $\sqrt{\dot{\phi}_k}(e_1 \circ \phi_k, f_1 \circ \phi_k)$
- 2) Compute the vertical component

$$\begin{aligned} (v_e, v_f) = & \left(\frac{1}{2} \alpha_\theta e_{\phi_k}(1) + \alpha \frac{d}{d\theta} e_{\phi_k}(1), \frac{1}{2} \alpha_\theta f_{\phi_k}(1) + \alpha \frac{d}{d\theta} f_{\phi_k}(1) \right) \\ & \text{of } (e'_{\phi_k}(1), f'_{\phi_k}(1)) \text{ where } \alpha : \mathbb{S}^1 \rightarrow \mathbb{R} \text{ solves} \\ & \Omega_\theta(e_{\phi_k}(1), e'_{\phi_k}(1) - v_e) + \Omega_\theta(f_{\phi_k}(1), f'_{\phi_k}(1) - v_f) = 0 \end{aligned} \quad (\text{II.23})$$

$$\text{and } \Omega_\theta(a, b) = a \cdot b_\theta - b \cdot a_\theta.$$

- 3) Set $\phi_{k+1} = \phi_k - \varepsilon \alpha$ where $\varepsilon > 0$ is small

Note that $\phi_k \in \text{Diff}(\mathbb{S}^1)$ and ϕ_0 is set to the identity on \mathbb{S}^1 . The idea of this algorithm is illustrated in Figure 1. The geodesic $(e_{\phi_k}(t), f_{\phi_k}(t)), t \in [0, 1]$ (in M) for large k will approximate the geodesic in B .

Figure 2 shows an example geodesic in B .

III. DYNAMICAL MODEL FOR DEFORMING SHAPES

We start by considering a simple “constant-velocity plus perturbation” model for a point moving in \mathbb{R}^n

$$\begin{aligned} \mu_k &= \mu_{k-1} + \nu_{k-1} \\ \nu_k &= \nu_{k-1} + \eta_{k-1} \end{aligned} \quad (\text{III.1})$$

where the state is $x_k = (\mu_k, \nu_k)$, η_{k-1} is a noise process, and μ represents the position and ν the velocity. When $\{\eta_k\}$ is a white Gaussian process, this is a discrete-time Brownian motion, or first-order random walk. This model can be considered to be a *first-order* approximation to a more complicated dynamical system when the temporal statistics

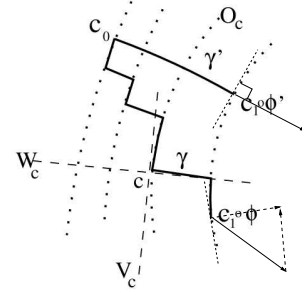


Fig. 1. The dashed lines represent equivalence classes of curves (O_c is the orbit $[c]$), and W_c is the horizontal space to c . To compute geodesics in B , we compute the geodesic in M between c_0 and $c_1 \circ \phi$ (the staircase path), project $\dot{\gamma}$ to its vertical component (tangent to $O_{c_1 \circ \phi}$), and move c_1 to another representative determined by the vertical component, and iterate the process until the vertical component becomes zero.

of $\{\eta_k\}$ are colored. We assume that we are given noisy measurements of the first component of the state, i.e.,

$$y_k = \mu_k + \xi_k \quad (\text{III.2})$$

where ξ_k is the measurement noise.

We now generalize the above dynamical model from \mathbb{R}^n to the case of curves. Denote with $\mu_k \in B$ the *deforming contour*, and $\nu_k \in T_{\mu_k}B$ its *velocity* at time k . The state at time k is $x_k = (\mu_k, \nu_k)$. Note that we may define the analogous operation to addition, i.e., $\mu_k + \nu_k$ in a Riemannian space, by using the exponential map. Also, since ν_k and ν_{k-1} are not in the same space in the case of curves (i.e., $\nu_k \in T_{\mu_k}B$ and $\nu_{k-1} \in T_{\mu_{k-1}}B$), the expression $\nu_k = \nu_{k-1} + \eta_{k-1}$ is not defined, and we must transport ν_{k-1} to $T_{\mu_k}B$ via parallel transport. Therefore, “the constant-velocity” model in the space of curves becomes

Definition III.3 (Discrete Brownian Motion on Curves):

$$\mu_k = \exp_{\mu_{k-1}}(\nu_{k-1}) \in B \quad (\text{III.3.*})$$

$$\nu_k = P_{\mu, k-1, k}(\nu_{k-1} + \eta_{k-1}) \in T_{\mu_k}B \quad (\text{III.3.†})$$

where $x_k = (\mu_k, \nu_k) \in TB$ is the state, $\eta_{k-1} \in T_{\mu_{k-1}}B$ is a *noise process*, and $P_{\mu, k-1, k}$ denotes parallel transport along the geodesic connecting μ_{k-1} to μ_k . Note that the noise process lives in a linear space, where it is easy to define a Gaussian distribution.

We will assume that noisy samples of the contour μ_k are available at each time k , for instance from any segmentation scheme from the active contour literature:

Definition III.4 (Measurement Model):

$$y_k = \exp_{\mu_k}(\xi_k) \in B \quad (\text{III.4.*})$$

where $\xi_k \in T_{\mu_k}B$ is the measurement noise.

The measurement is a noisy version of the first component of the state, μ_k . Again, notice that ξ_k lives on a linear space, where a Gaussian distribution can be easily defined.

IV. FILTERING DEFORMING SHAPES

In this section, the goal is to devise a recursive estimation procedure to estimate the state of the dynamical system, (μ_k, ν_k) , i.e., the shape and velocity of a moving object,

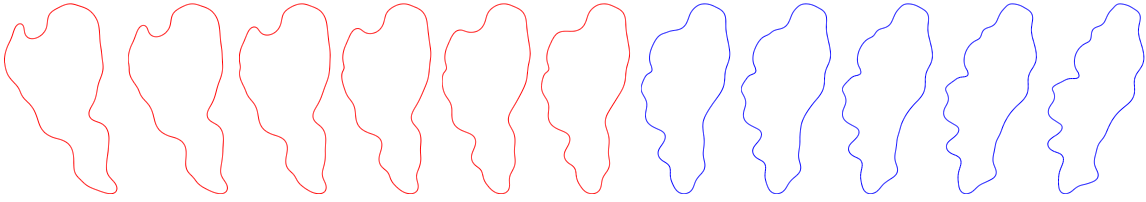


Fig. 2. Example illustration of a geodesic in B . The geodesic is computed between the two red end curves (time $t = 0$ and $t = 1$), and intermediate curves (interpolation) are between in red. The blue curves ($t = 1$ to $t = 2$) are the continuation of the geodesic (extrapolation) from the last red curve.

introduced in the previous section from measurements y_k obtained from the time-varying image, I_k . We start by reviewing the classical linear finite-dimensional (Luenberger) observer in \mathbb{R}^n , then generalize it to the space of curves.

An observer in \mathbb{R}^n for the dynamical system (III.1) and measurement model (III.2) takes the form

$$\hat{\mu}_{k|k-1} = \hat{\mu}_{k-1|k-1} + \hat{\nu}_{k-1|k-1} \quad (\text{IV.1})$$

$$\hat{\nu}_{k|k-1} = \hat{\nu}_{k-1|k-1} \quad (\text{IV.2})$$

$$\hat{\mu}_{k|k} = \hat{\mu}_{k|k-1} + K_\mu(y_k - \hat{\mu}_{k|k-1}) \quad (\text{IV.3})$$

$$\hat{\nu}_{k|k} = \hat{\nu}_{k|k-1} + K_\nu(y_k - \hat{\mu}_{k|k-1}) \quad (\text{IV.4})$$

the first two equations are the *state prediction*, the last two equations are the *update*, and $K_\mu, K_\nu > 0$ are called the *gains*, which can be chosen to satisfy some optimality criterion. The weakest requirement, for time-invariant models, is that the error

$$e_k = x_k - \hat{x}_{k|k}$$

between the state estimate and the true state approaches zeros as $k \rightarrow +\infty$.

The analogous observer in the case of the dynamical system on the space of curves takes the following form:

Definition IV.5 (Curve Observer): The prediction is

$$\hat{\mu}_{k|k-1} = \exp_{\hat{\mu}_{k-1|k-1}}(\hat{\nu}_{k-1|k-1}) \quad (\text{IV.5.*})$$

$$\hat{\nu}_{k|k-1} = P_{\hat{\mu}_{k-1|k-1}, \hat{\mu}_{k|k-1}}(\hat{\nu}_{k-1|k-1}) \in T_{\hat{\mu}_{k|k-1}}B; \quad (\text{IV.5.†})$$

where $P_{\hat{\mu}_{k-1|k-1}, \hat{\mu}_{k|k-1}}$ denotes parallel transport along the geodesic from $\hat{\mu}_{k-1|k-1}$ to $\hat{\mu}_{k|k-1}$.² The update is

$$\hat{\mu}_{k|k} = \hat{\mu}_{k|k-1} \quad (\text{IV.5.‡})$$

$$\hat{\nu}_{k|k} = \hat{\nu}_{k|k-1} + K \log(\hat{\mu}_{k|k-1}, y_k). \quad (\text{IV.5.§})$$

where $K > 0$ is a constant that can be chosen to trade off asymptotic tracking error with convergence speed. Note that the sum in the last equation is well defined as $\hat{\nu}$ lives on a linear space.

Remark IV.6: The Luenberger observer structure involves a direct effect of the measurements through a gain matrix K that has two components, K_μ affecting μ , and K_ν affecting ν . In this case the update equations would be $\hat{\mu}_{k|k} = \exp_{\hat{\mu}_{k|k-1}}(K_\mu \log(\hat{\mu}_{k|k-1}, y_k))$, and $\hat{\nu}_{k|k} =$

²Note that the parallel transport $P_{\hat{\mu}_{k-1|k-1}, \hat{\mu}_{k|k-1}}(\hat{\nu}_{k-1|k-1})$ is trivial to compute since it is the parallel transport along a geodesic of its own tangent vector. The parallel transport is computed as $\dot{\gamma}(1)$ where γ is the geodesic between $\gamma(0) = \hat{\mu}_{k-1|k-1}$ and $\gamma(1) = \hat{\mu}_{k|k-1}$, which is automatically obtained when evaluating (IV.5.*).

$P_{\hat{\mu}_{k|k-1}, \hat{\mu}_{k|k}}(\hat{\nu}_{k|k-1} + K_\nu \log(\hat{\mu}_{k|k-1}, y_k))$. The latter requires parallel-transport of a tangent vector that is not tangent to the geodesic path to be transported along, which entails solving a differential equation numerically. While this is certainly possible, we consider the simplified observer structure where the correction occurs at the velocity level, and therefore $K_\mu = 0$.

V. EXPERIMENTS

The following experiments are meant to illustrate the behavior of the dynamical model that we have constructed, and thus, we have chosen a very basic segmentation technique to obtain the measurements y_k . The measurements y_k are obtained by performing an active contour segmentation using the Chan-Vese energy [4] and Sobolev active contours [22], [23] with the initialization being the state contour prediction, $\hat{\mu}_{k|k-1}$. The initialization of the state contour, $\hat{\mu}_{0|0}$, is chosen by hand and the state velocity $\hat{\nu}_{0|0}$ is chosen to be zero. Further, we have chosen the gain $K = 0.2$ unless specified otherwise. The red curve indicates the state prediction contour $\hat{\mu}_{k|k-1}$, the blue arrows indicate the state prediction velocity $\hat{\nu}_{k|k-1}$, and the green curves indicate the measurement y_k all at frame k .

In the first experiment (Fig. 3), we track a circle that continuously deforms (by a non-affine deformation) into two joined blobs. The data is corrupted by a full occlusion in frames 6-11 (the sequence ranges from 1-13). In frame 1, we choose the contour initialization to match the circle's boundary. In the top row, we have used a dynamical model and filter on the affine motion parameters of the object as is typical in prior work. In the bottom row, we have used the proposed method, which defines a dynamical model and filters on the space of curves. At the moment of occlusion ($t = 6$), we set the gain $K = 0$ in which case the filter ignores the measurements y_k , and moves according to state dynamics for $t \geq 6$ with the initial velocity $\hat{\nu}_{6|6}$. In the case that only affine dynamics are considered, the shape of the object is not captured accurately. The dynamical model on arbitrary deformations more accurately captures the deformation and shape of the object.

In Fig. 4 and 5, we track a deforming flatworm in the ocean. Fig. 4 shows the proposed filtering technique applied to the sequence. The experiment demonstrates that the constant velocity plus perturbation model (whose trajectory is shown in red and blue arrows) does a good job of predicting and extrapolating the boundary and motion of the object. In Fig. 4, we compare our proposed model to a filtering

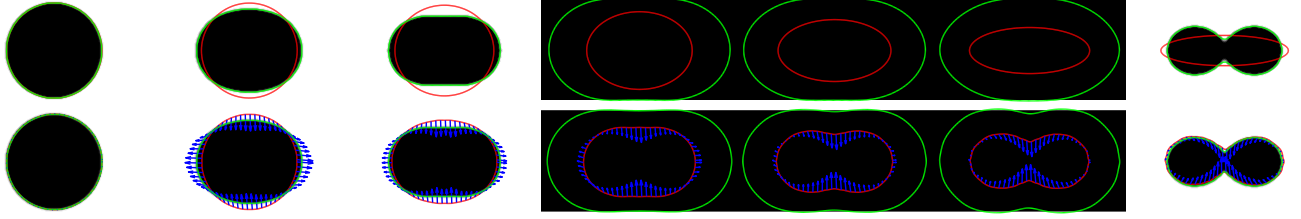


Fig. 3. Tracking a synthetic deforming circle through a total occlusion: this experiment demonstrates the need for the dynamical model that extrapolates the shape. In the first few frames, where there is no occlusion, the image segmentation (green) alone provides the correct tracking. However, when the occlusion appears, the image segmentation is wrong, but the dynamical model extrapolates the shape (red) and velocity (blue) of the contour (middle frames). A dynamical model with only affine motion dynamics (top row) does not capture the deformation of this object through the dynamics. The last frame shows that the state dynamics have accurately captured the object. Red: $\hat{\mu}_{k|k-1}$, the blue: $\hat{\nu}_{k|k-1}$ and the green: y_k .

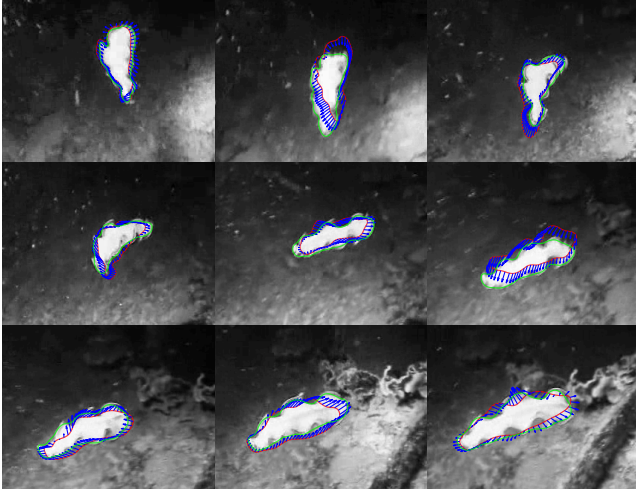


Fig. 4. Tracking a flatworm (left to right, top to bottom) using the proposed filtering technique: the red curve is $\hat{\mu}_{k|k-1}$, the blue arrows are $\hat{\nu}_{k|k-1}$, and the green curve is the measurement y_k . This experiment demonstrates the dynamics of the contour and deformation under the constant velocity plus perturbation model, which does a decent job of modeling the dynamics of the flatworm.

strategy that only filters and models on the affine motion. As can be seen from the figure, the proposed model which filters arbitrary deformations, gives far greater accuracy of the deforming object than the affine motion model. Indeed, the proposed model predicts a contour that is close enough to the desired local minimum of the Chan-Vese energy that an accurate measurement is obtained. In contrast, the affine model predicts a contour that is far away from the desired local minimum that the measurement leaks into light portions of the background.

VI. CONCLUSION

We have introduced what is, to the best of our knowledge, the first ever filtering and prediction scheme on the infinite-dimensional space of shapes, defined as simple, closed planar contours undergoing general diffeomorphisms. Previous work has either attempted to “separate” the “motion” (a finite-dimensional group) from the “deformation”, and defined observers for dynamical models on the finite-dimensional motion parameters, or has restricted the set of allowable deformations to finitely-parametrized classes, for

instance obtained from manually obtained training data. The problem with the former approach is that it fails to predict deformations; as an object undergoes an occlusion, the tracker can extrapolate its affine motion, but not its deformation. We have shown that predicting deformations allows us to significantly decrease prediction error. The problem with the latter approach is that it requires having training data available for the classes of objects and deformations that one wishes to track. While this is realistic for objects like humans walking, it becomes prohibitive when one wants to consider more gaits (limping, running, hopping), or more objects (flatworms, jellyfish, hurricanes) for which training data may not be available.

Deriving a dynamic observer on the space of curves is no easy feat, and we have had to tap onto the most recent advances in the shape analysis and active contour literature, classical results in Differential and Riemannian geometry, and classical results in prediction and filtering theory. We have illustrated the case of (first-order) random walk dynamics, but our approach can be easily extended to any linear dynamics, for instance auto-regressive moving-average models of any order. This is made possible by the fact that the stochastic processes driving the dynamics are defined on the tangent spaces to the state-space, which are linear and therefore standard tools from systems theory can be applied, albeit with care because these linear spaces are still infinite-dimensional.

While one may wish to bypass the significant mathematical burden by discretizing the objects of interest at the outset, for instance by using a piecewise linear contour, or a spline or Bezier curve, this introduces difficulties later on. In fact, the location of control points or vertices can move while keeping the data unchanged, which results in an un-observable model, and therefore cause spurious dynamics in the observer. This problem is further exacerbated when one wishes to generalize the approach to tracking surfaces in space. Our approach avoids these representational issues by modeling directly the native objects – closed simple planar contours – in the space where they belong, leaving the discretization to the last stage of computation, which is the numerical integration of the partial differential equations implementing the observer. Our approach has been demonstrated on sequences of deforming objects.

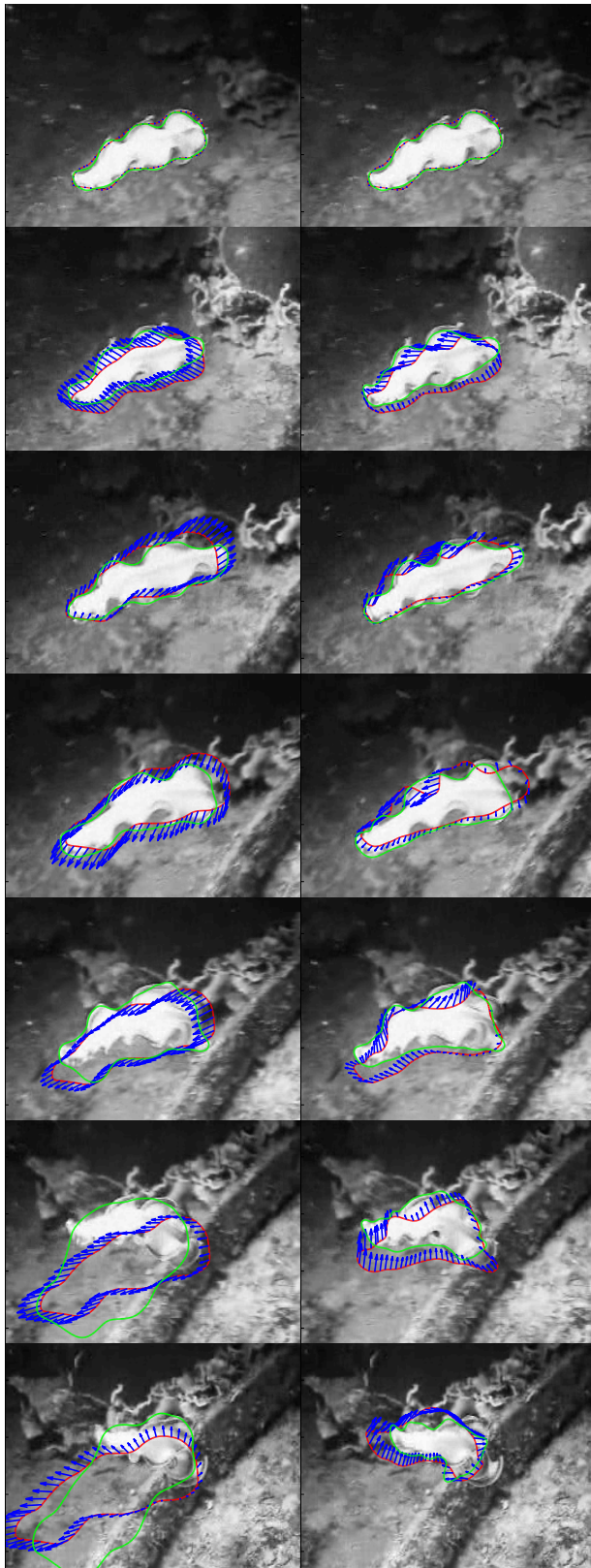


Fig. 5. Comparison between dynamical model only on the motion (scale and translation) on the left column and dynamical model on both the motion and deformation (right column). On the left column it can be seen that the predicted shape fails to adapt to the newly deformed object; on the right column, where both motion and deformation are extrapolated, the object is predicted with far greater accuracy. Red: $\hat{\mu}_{k|k-1}$, blue: $\hat{\nu}_{k|k-1}$, and green: y_k .

ACKNOWLEDGEMENTS

This research was funded by AFOSR FA9550-09-1-0427.

REFERENCES

- [1] M. F. Beg, M. I. Miller, A. Trounev, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.
- [2] A. Blake and R. Brockett. On snakes and estimation theory. In *IEEE CDC*, 1994.
- [3] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proceedings of the IEEE Int. Conf. on Computer Vision*, pages 694–699, Cambridge, MA, USA, June 1995.
- [4] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, February 2001.
- [5] G. Charpiat, O. D. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.
- [6] D. Cremers. Nonlinear dynamical shape priors for level set segmentation. In *CVPR*. IEEE Computer Society, 2007.
- [7] D. Cremers and S. Soatto. Motion competition: A variational approach to piecewise parametric motion segmentation. *International Journal of Computer Vision*, 62(3):249–265, 2004.
- [8] M. Delfour and J. P. Zolésio. Shape identification via metrics constructed from the oriented distance function. *Control and Cybernetics*, 34(1):137–164, 2005.
- [9] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20:303–353, 1998.
- [10] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *IJCV*, 1(29):5–28, 1998.
- [11] J. Jackson, A. Yezzi, and S. Soatto. Tracking deformable moving objects under severe occlusions. In *IEEE Conference on Decision and Control*, Dec. 2004.
- [12] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331, 1987.
- [13] R. Mathias. Evaluating the frechet derivative of the matrix exponential. *Numerische Mathematik*, 63:213–226, 1992.
- [14] P. Michor and D. Mumford. Riemannian geometries on the space of plane curves. *J. Eur. Math. Soc.*, 8:1–48, 2006.
- [15] P. W. Michor and D. Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23:76–113, 2007.
- [16] W. Mio and A. Srivastava. Elastic-string models for representation and analysis of planar shapes. In *CVPR (2)*, pages 10–15, 2004.
- [17] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
- [18] M. Niethammer, P. A. Vela, and A. Tannenbaum. Geometric observers for dynamically evolving curves. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(6):1093–1108, 2008.
- [19] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on the Hamilton-Jacobi equations. *J. Comp. Physics*, 79:12–49, 1988.
- [20] N. Paragios and R. Deriche. Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision. *International Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics*, 13(2):249–268, June 2002.
- [21] Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi. Particle filtering for geometric active contours and application to tracking deforming objects. In *IEEE CVPR*, 2005.
- [22] G. Sundaramoorthi, A. J. Yezzi, and A. Mennucci. Sobolev Active Contours. *International Journal of Computer Vision*, 73(3):345–366, 2007.
- [23] G. Sundaramoorthi, A. J. Yezzi, and A. Mennucci. Coarse-to-fine segmentation and tracking using Sobolev Active Contours. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(5):851–864, 2008.
- [24] D. Terzopoulos and R. Szeliski. *Active Vision*, chapter Tracking with Kalman Snakes. MIT Press, 1992.
- [25] N. Vaswani, Y. Rathi, A. Yezzi, and A. Tannenbaum. Pf-mt with an interpolation effective basis for tracking local contour deformations. *IEEE Trans. on Image Processing*, 2008.
- [26] C. Wasshuber. *Computational Single-Electronics*. Springer, 2001.
- [27] L. Younes. Computable elastic distances between shapes. *SIAM J. Appl. Math.*, 58(2):565–586, 1998.
- [28] L. Younes, P. W. Michor, J. Shah, and D. Mumford. A metric on shape space with explicit geodesics. *Rend. Lincei Mat. Appl.*, 9:25–57, 2008.